

GLUE Schema Status, open issues and proposed solutions

FermiLab Meeting
02 Oct 2003

Sergio Andreozzi
INFN-CNAF
Bologna (Italy)

sergio.andreozzi@cnaf.infn.it



OUTLINE

- I. GLUE Schema
 - I. Short Intro
 - II. Current status
- II. Revision under discussion for
 - I. Generic Grid Service
 - II. Storage
 - III. Computing
- III. Extension for monitoring
 - I. GridICE and the GLUE Schema

GLUE Schema - Short history

- ₁ Part of the GLUE collaboration
- ₁ Promoted by DataTAG (EU) and iVDGL (US)
- ₁ Contributions from DataGrid, Globus, PPDG and GriPhyn

- ₁ GLUE Schema activity is started in April 2002

A common information model for Grid resources is one of the main tasks

GLUE Schema: modeling guidelines

- F Clear separation between system and service entities
 - § System: a set of connected items or devices which operate together as a functional whole
 - § Service: actions that form a coherent whole from the point of view of service providers and service requesters
- F Generalization
 - § capture common aspects for different entities providing the same functionality (e.g., uniform view over different batch services)
- F Deal with both monitoring needs and discovery needs
 - § Monitoring: concerns those attributes that are meaningful to describe the status of resources (e.g., useful to detect fault situation)
 - § Discovery: concerns those attributes that are meaningful for locate resources on the base of a set of preferences/constraints (e.g., useful during matchmaking process)

GLUE Schema: status

- F Current Conceptual model - version 1.1
 - Finalized in Mar '03
 - Some fix to Computing/Storage resources model from v1.0
 - Model of network resources

GLUE: Implementation Status

¹ Implementation status - version 1.1

- q For Globus MDS 2.x (part of GT 2.x):
 - LDAP Schema (DataTAG WP4)
 - Info providers both computing (EDG WP4, valid for PBS, LSF and Condor) and storage resources (EDG WP5, valid for trivial file system and edg-se)
- q For EDG R-GMA:
 - Relational schema (EDG WP3)
 - Info providers for computing and storage resources translate output of LDAP info provider in a suitable format to be stored in the relational model (EDG WP3)
 - Info providers for network resources (EDG WP7+DataTAG WP4)
- q For Globus MDS 3.x (part of GT 3):
 - XML Schema for computing resources (Globus)
 - Info provider (Globus)

GLUE Schema: Deployment Status

¹ Included in:

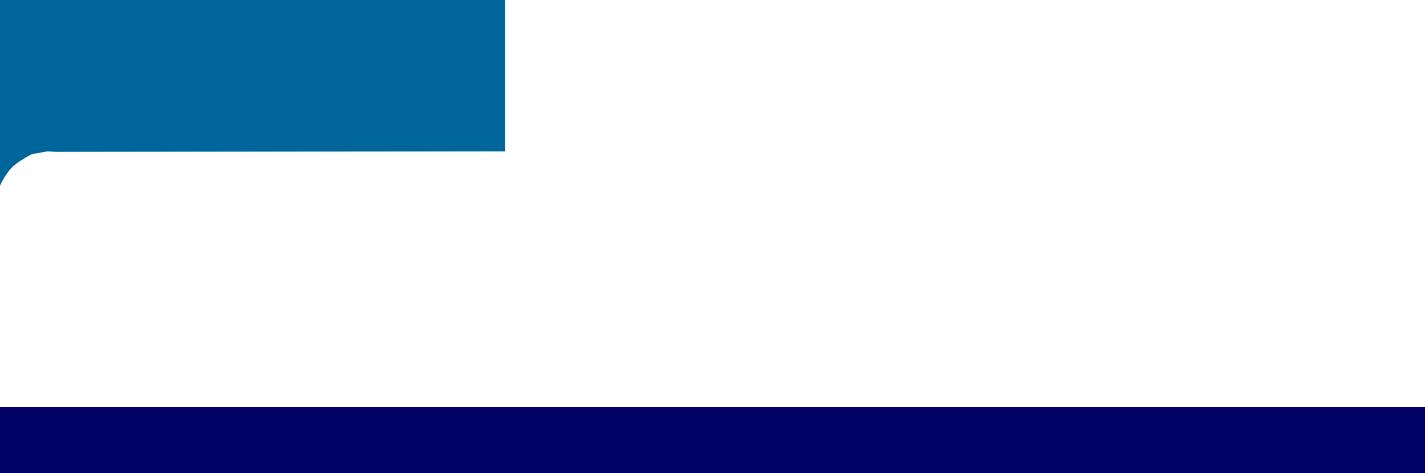
- **DataGrid 2.0**
 - ¹ with mixed R-GMA/MDS2 scenario
- **VDT 1.1.6 and later (MDS2)**
- **LCG0 (MDS2)**
- **LCG1 (MDS2)**
- **Globus Toolkit 2.x**
 - ¹ as optional, only for computing resources
- **Globus Toolkit 3**
 - ¹ as optional, only for computing resources

Important notice 1

- 1 All the following proposals have been discussed in different contexts, but never brought officially within the GLUE Schema Group
- 1 Since every schema change can have sensitive impact on the running software, we want a mature proposal

Important notice 2

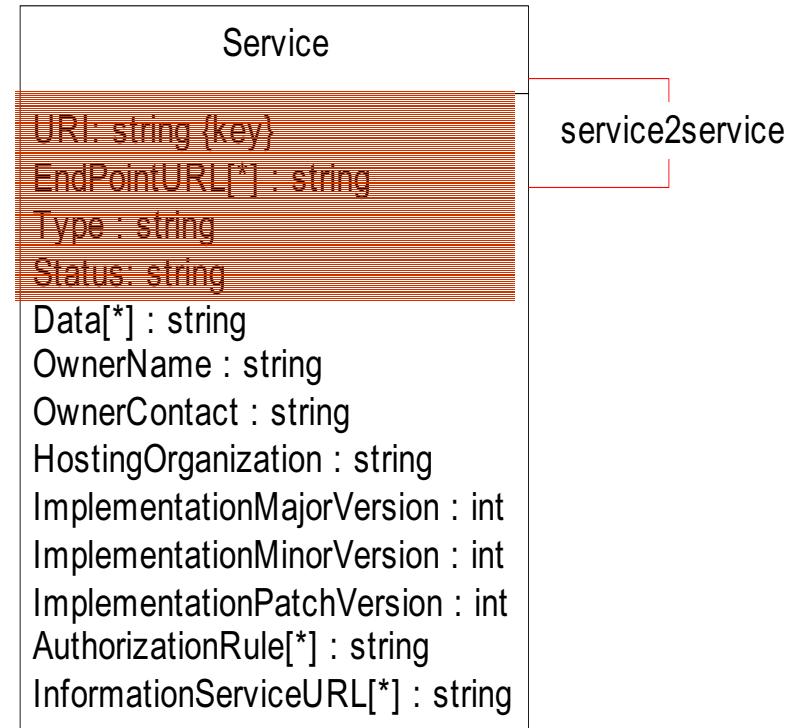
- 1 It is desirable that all the proposed modification will be divided in two sets:
 - Short term/urgent modification (to be discussed and deployed before the end of the year)
 - Medium term modification



Discussion about a generic Service model

GLUE Schema - defining a generic service schema

- 1 **URI**: unique identifier for the service
- 1 **EndPointURL**: set of service access points that may correspond to different network addresses and bindings (e.g., secure/unsecure ports, multiple gatekeepers to a batch queue)
- 1 **Type**: the service type, we should define an enumeration (e.g. edg-replica-optimization, edg-replica-metadata-catalog)
- 1 **Status**: the service status



GLUE Schema - defining a generic service schema

- 1 **Data:** set of generic attributes
- 1 **OwnerName:** The name of the primary owner for the service, if one is defined. The primary owner is the initial support contact for the Service
- 1 **OwnerContact:** A string that provides information on how the primary owner of the Service can be reached (e.g. phone number, email address, ...)
- 1 **HostingOrganization:** the organization id/name that hosts the service

Service
URI: string {key}
EndPointURL[*] : string
Type : string
Status: string
Data[*] : string
OwnerName : string
OwnerContact : string
HostingOrganization : string
ImplementationMajorVersion : int
ImplementationMinorVersion : int
ImplementationPatchVersion : int
AuthorizationRule[*] : string
InformationServiceURL[*] : string

service2service

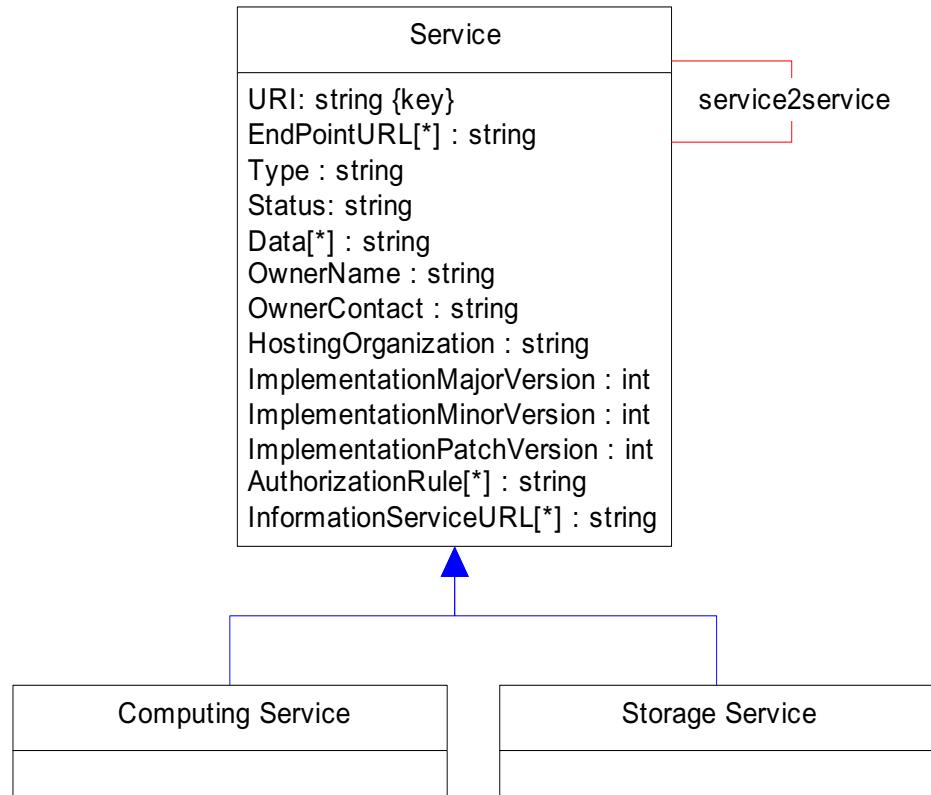
GLUE Schema - defining a generic service schema

- 1 **Major/Minor/PatchVersion:**
service implementation
major/minor/patch version
- 1 **AuthorizationRule[*]:** set of
strings expressing authorization
rules for this service
- 1 **InformationServiceURL[*]:** URL
of info service that can provide
for info about the service (e.g. a
GRIS, an R-GMA producer)
 - Q1. is it reasonable to have
multiple instances of this?
 - Q2. can we use it also for
pointing at specific monitoring
services?

Service
URI: string {key}
EndPointURL[*] : string
Type : string
Status: string
Data[*] : string
OwnerName : string
OwnerContact : string
HostingOrganization : string
ImplementationMajorVersion : int
ImplementationMinorVersion : int
ImplementationPatchVersion : int
AuthorizationRule[*] : string
InformationServiceURL[*] : string

GLUE Schema - defining a generic service schema

- 1 Once this is defined, I'd like to see it as a generalization of all grid services
- 1 already available (e.g., computing and storage) can inherit from it and add/redfine service specific aspects



Discussion about Storage Resources revision

GLUE Storage Service/Space

1 Storage Service:

- grid service **identified by a URI**
- it **manages storage extents** in term of Storage Spaces
- it has a **type** (e.g., edg-se, fs-only, srm)
- it offers a **set of data access protocols** to access files in the Storage Spaces or transfer them in/out

GLUE Storage Service/Space

- 1 **Storage Space:** portion of a logical storage extent identified by a local unique identifier that:
 - is assigned to a Virtual Organization
 - It may be associated to a directory of the underlying file system (e.g., /permanent/CMS); this association may be many-to-one
 - has a set of policies (MaxFileSize, MaxData, MaxNumFiles, MaxPinDuration, Quota)
 - has a set of access control base rules (e.g., to privilege some VO user against some other)
 - has a state (e.g., available space)

GLUE Storage Space - Open issue missing Unique ID

- 1 **ISSUE:** missing Unique Identifier (there was the wrong assumption of a 1-to-1 relationship with the directory)
- 1 **SOLUTION:** add GlueSAUniqueID
 - The LDAP DIT will change as follow:
 - ↳ dn: GlueSARoot=...,GlueSEUniqueID=...,mds-vo-name=local, o=grid
 - dn:GlueSAUniqueID=..., GlueSEUniqueID=...,mds-vo-name=local,o=grid
 - and GlueSARoot will be just an attribute not part of the key

GLUE Storage Space - Open issue free space attribute computation

- 1 **ISSUE:** free space attributes are per Storage Space;
 - schema issue: how do I distinguish if the free space is exclusive or shared?
- 1 **SOLUTION:**
 - Split Quota attribute in BestEffortQuota and GuaranteedQuota,
 - Then couple the value of BestEffort/Guaranteed Quota with the Free/Used space, e.g.:



GLUE Storage Space - Open issue missing state attribute

- 1 **ISSUE:** there is no state attribute for an SE

- 1 **SOLUTION:**
 - Add GlueSEStatus

GLUE Storage Library - Open issue what do we need?

1 ISSUE:

- The Storage Library is still a raw concept;
- Referring to the storage system, the experience says that what need is to represent the state of the machine where the service runs

1 SOLUTION:

- Replace the Storage Library with the Host, modeling that it is an access machine to a storage system (it can be done using the HostRole modeled in the extensions for monitoring)

GLUE Data Access Protocol - Open issue unique id for a data access prot.

- 1 **ISSUE:** at present a data access protocol is identified by its type;
 - for scalability issues, the same protocol can be used from multiple ports, or from multiple machines
- 1 **SOLUTION:**
 - Add AccessProtocol.URI, this is the key
 - 1 E.g., gsiftp://hostname:port/base_dir
 - Remove AccessProtocol.port (not needed anymore, since it is part of the URI)

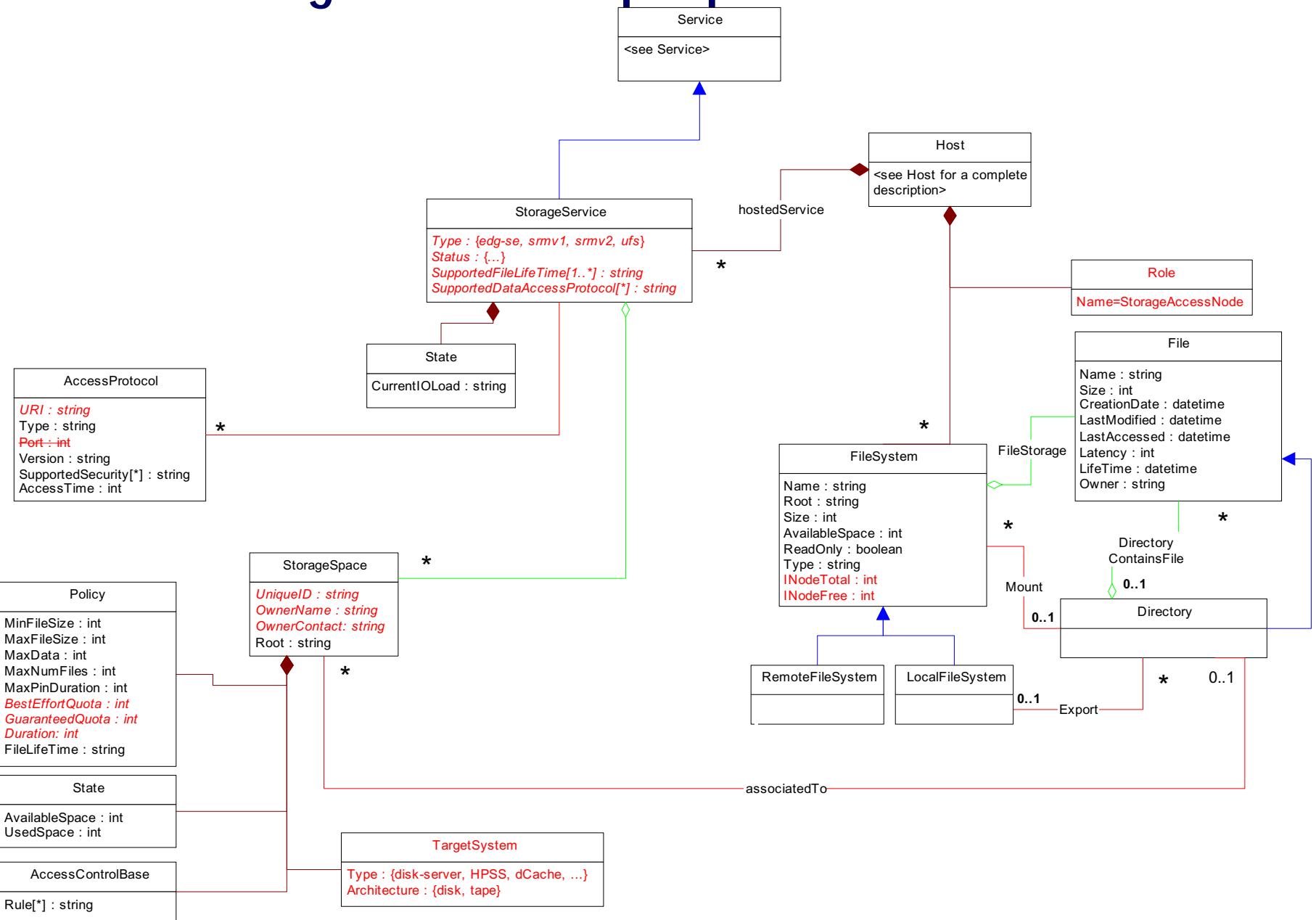
GLUE CE-SE Bind - Open issue mounted dir/exported dir

- 1 **ISSUE:** at present, there is the assumption that the CE mounted dir name is equal to the exported SE dir name; this is a limitation because if two siteAdmins of two different SEs use the same name for the exported dir, then they cannot be mounted on the same cluster
- 1 **POSSIBLE SOLUTIONS:**
 - 1. we don't make any assumption on the SE exported dir name, the CESEBind.Accesspoint is the local mount dir
 - 2. if we need the name of the exported dir, we add one more attribute, e.g., CESEBind.ExportedDir
- 1 **MORE QUESTIONS:**
 - Does this relationship refer to a CE-SE couple or CE-SA (SA=Storage spAce)?

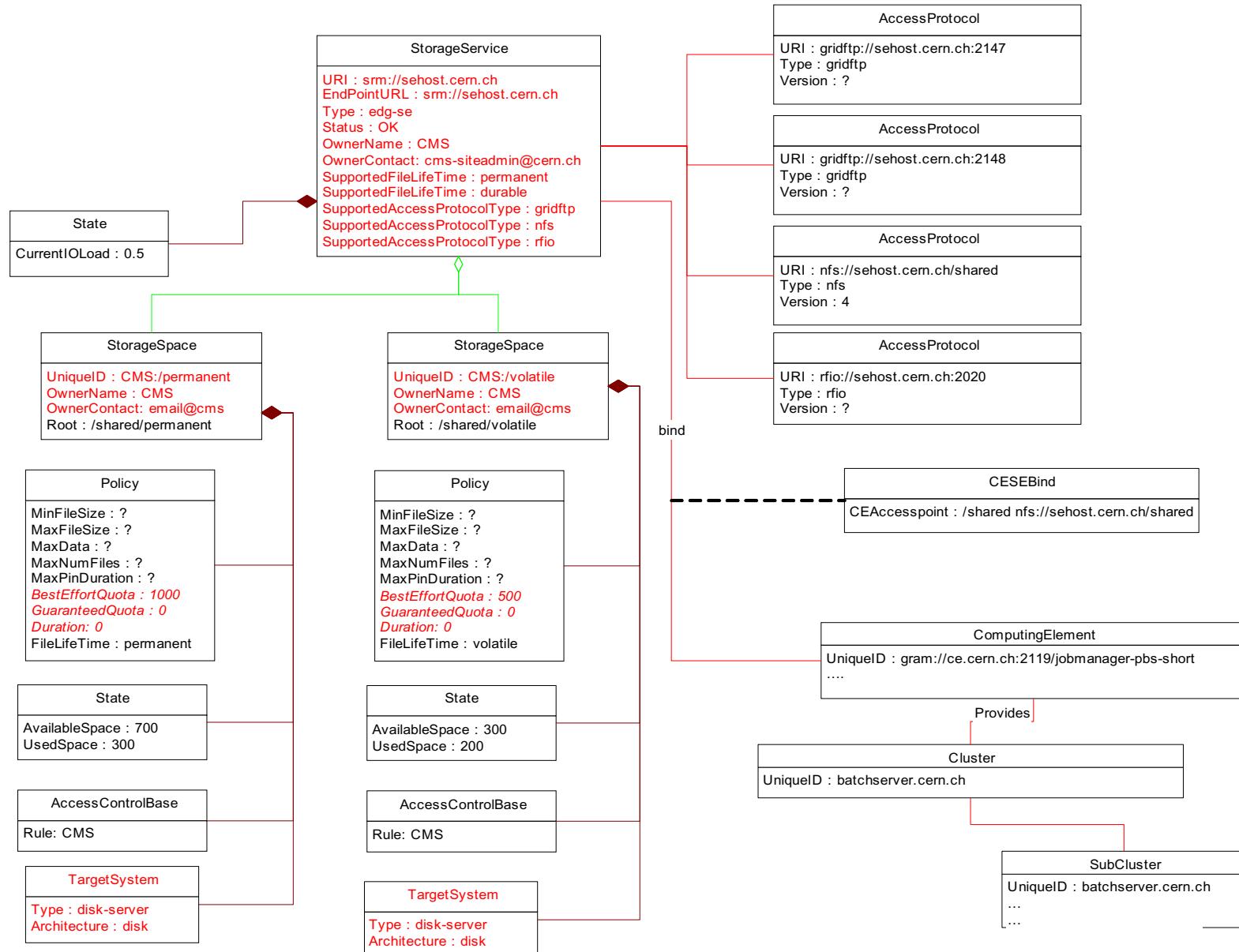
GLUE SE - Open issue Current IO Load

- 1 **ISSUE:** at present, there is an input/output load attribute for the storage service, but no good metric has been defined
- 1 **SOLUTION:**
 - ? Any proposal so that loads from different storage service can be meaningfully compared ?

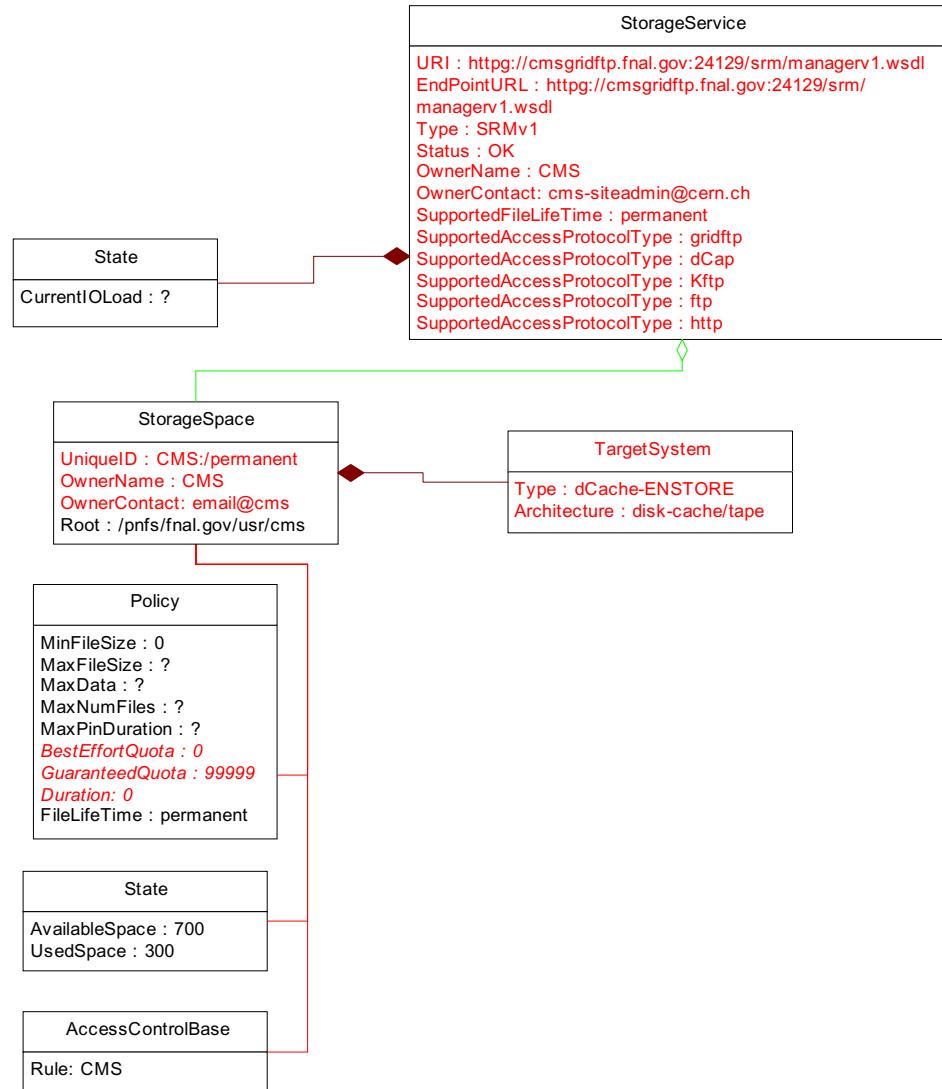
Storage Service - proposal revision



USE CASE - edg-se (to be checked)



USE CASE - SRMv1@FermiLab



Thanks to Michael Ernst and Timur Perelmutov

GLUE CE - Open issue from Computing Element to Computing Service

- The 'Element' usually reminds to physical entities
- Computing Element or CE is frequently associated to the access node machine where the gatekeeper is running
- In our context, we can speak about 'Computing Service'
- My proposed definition:
- **Computing Service**: uniquely identified grid service that can provide a user software application for computing power in a certain execution environment
- Any comments/suggestions?

GLUE CE - Open issues

Free/Total CPU concept

1 ISSUE (by example):

- I have three worker nodes dual-processor managed by LSF;
- I configure one queue and I assign all $2 \times 3 = 6$ CPUs, each CPU can run three jobs (MaxRunJobs=3*6=18);

CPU1	CPU2
1	2
3	2
0	0

X=RunJobs

CE.Info.TotalCPUs=?

Do we refer to physCPU
or to job slots?

CE.State.FreeCPUs=?

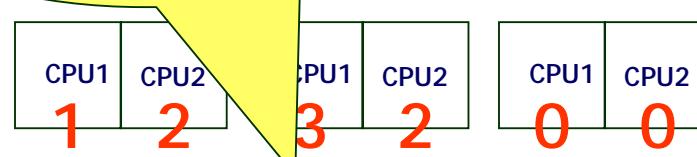
GLUE CE - Open issues

Free/Total CPU concept

1

ISSUE (b)

- I have a solution for this problem. With this solution you can have information about both physical CPUs and job slots
- I have a solution for this problem. I have 3 physical processors managed by LSF. Each processor has 2*3=6 CPUs, each slot can run 3 jobs (RunJobs=3*6=18);



X=RunJobs

~~CE.Info.TotalCPUs~~

CE.State.FreeJobSlots=10

~~CE.State.FreeCPUs~~

CE.Policy.AssignedJobSlots=18

CE.Policy.AssignedCPUs=6

GLUE SubCluster - Open issue EDG Constraint - only one

1 ISSUE:

- The limitation is due to a compromise solution, one entity for two different concepts:
 - 1 How to describe in summary an homogenous set of hosts part of a cluster? (no relationships with queue)
 - 1 How to describe in summary a set of hosts assigned to a queue, so that the matchmaking process is efficient?

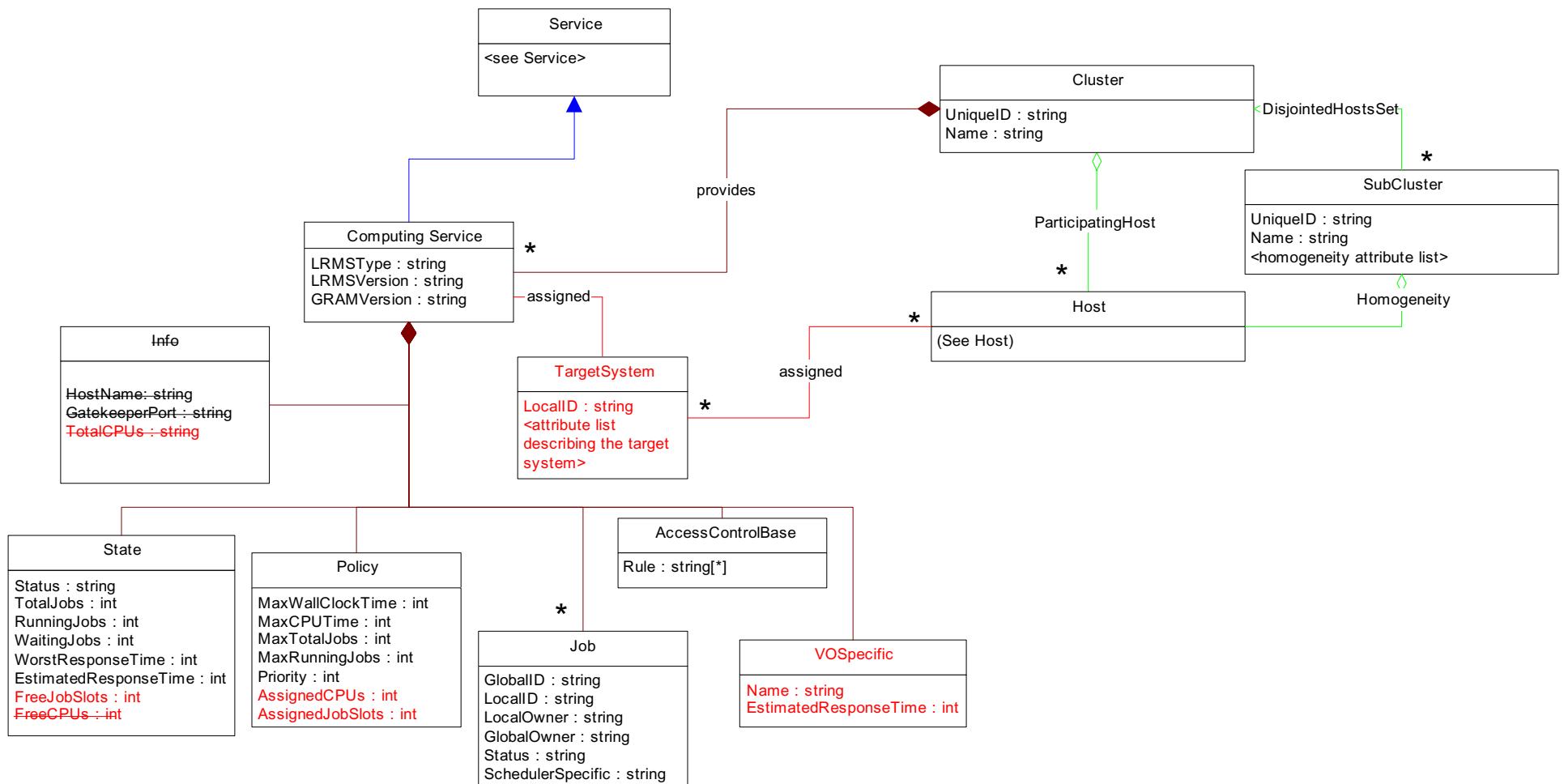
1 SOLUTION:

- Use an entity for each concept; need for introducing a new entity in relationships with both CE and hosts (given a CE, this describes in summary characteristics of assigned hosts)

1 NOTICE:

- the subcluster is in relationships with Cluster and Hosts, since it describes homogeneous set of hosts that hold the property of being part of a certain cluster

Computing Service - proposal revision

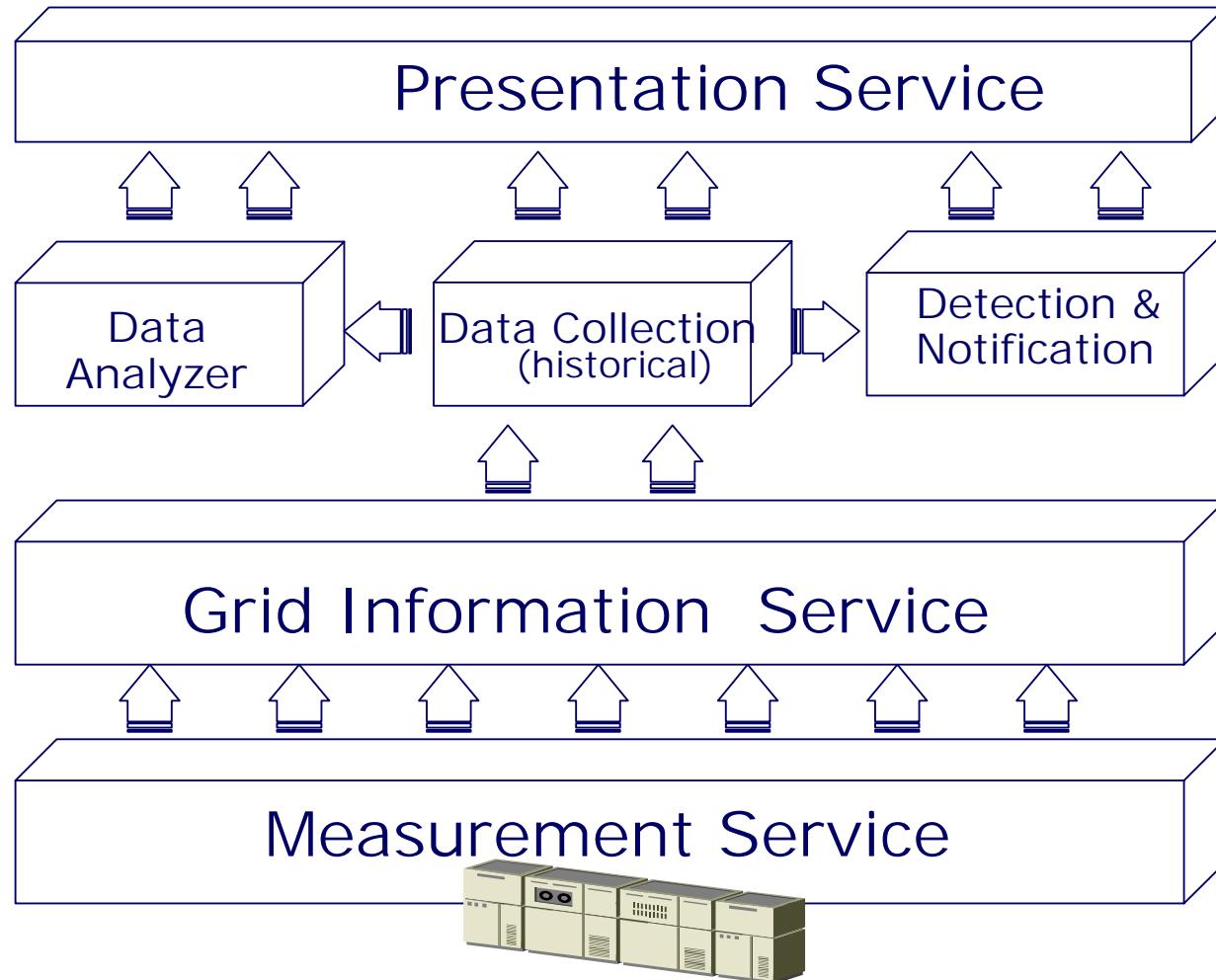


Discussion about GridICE monitoring extensions of the GLUE Schema

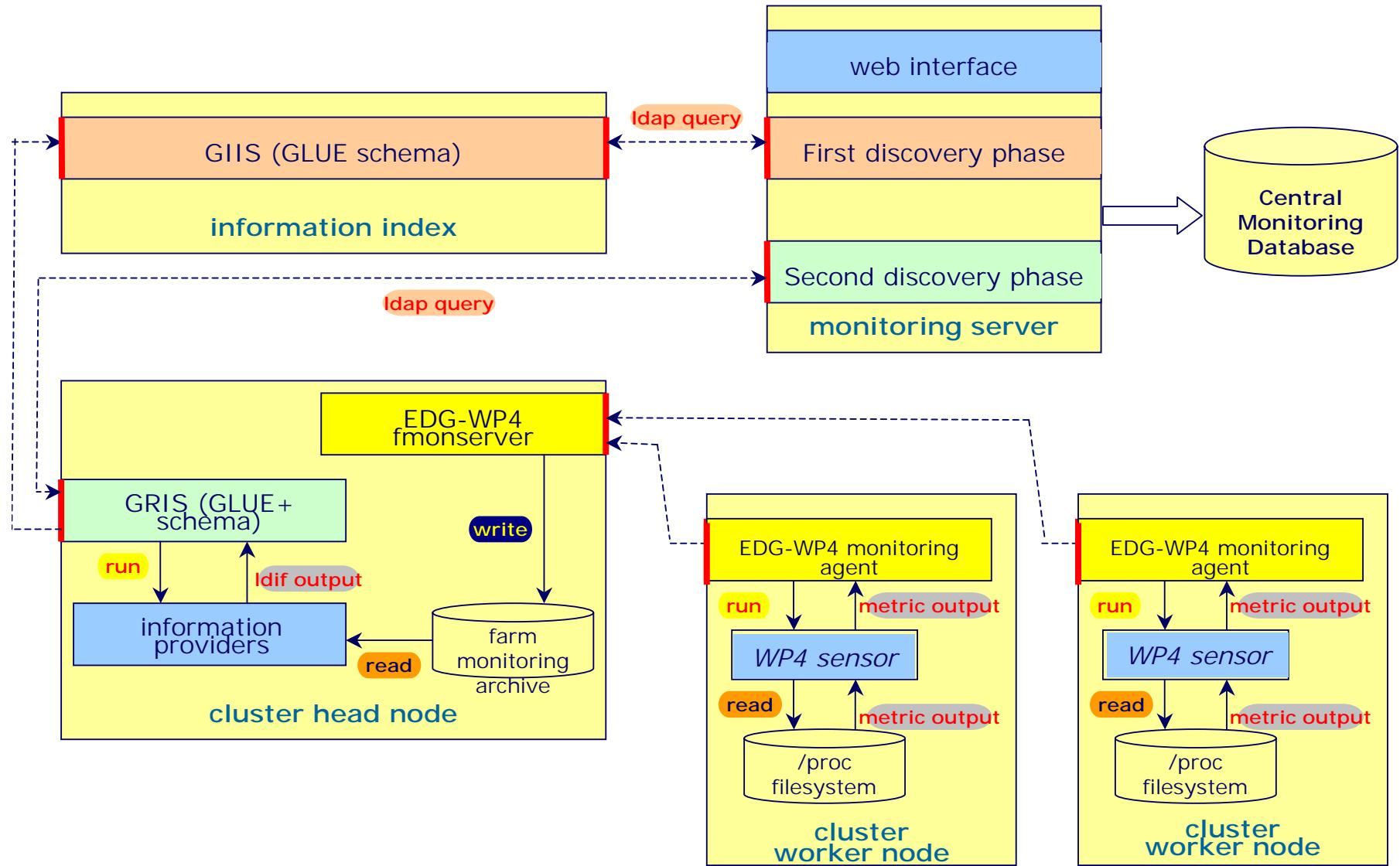
A couple of sentences about GridICE

- ₁ Grid Monitoring Service developed by INFN as part of the EU DataTAG project (WP4)
- ₁ Started on January 2003
- ₁ First release targeted at the current HEP Grid middlewares (EDG 1.4/2.0, LCG 0/LCG 1)
- ₁ Close collaboration with LCG project in order to meet requirements of a monitoring service for a Grid Operation Center

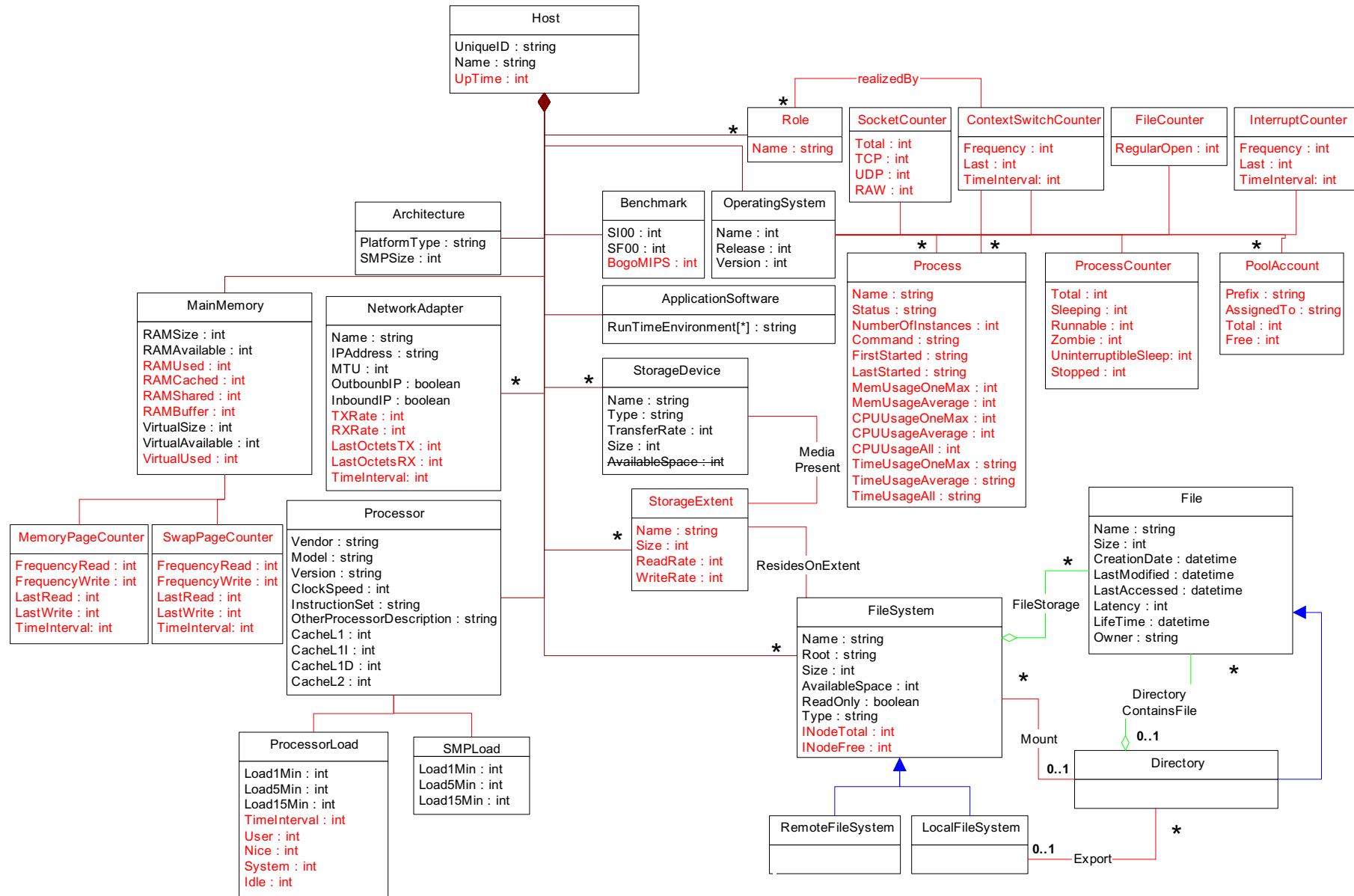
GridICE: data flow



GridICE functional view



Extensions to GLUE Schema (in red)



Towards a common extension of the GLUE Schema for monitoring

- ₁ DataTAG/LCG already have in place extensions to the GLUE Schema for monitoring purposes
- ₁ Grid3 project is defining other extensions

Towards a common extension of the GLUE Schema for monitoring

- ₁ We all would benefit from a common extension
- ₁ Resources from different grids (e.g., INFN-Grid, LCG-Grid and Grid3-Grid) will advertise monitoring data using a common schema
- ₁ The resource measurement service can be any, MDS will do the work of uniform access to data (e.g., within GridICE we currently interface EDG FMON, Grid3 is supporting Ganglia)

REFERENCES

- 1 GLUE Schema Official documents
<http://www.cnaf.infn.it/~sergio/datatag/glue>
- 1 S. Andreozzi, "GLUE Schema implementation for the LDAP model", Technical report, first draft, 29/05/03
<http://www.cnaf.infn.it/~sergio/publications/Glue4LDAP.pdf>
- 1 S. Andreozzi, "Working draft", (to be updated)
<http://www.cnaf.infn.it/~sergio/datatag/glue/working/SE>
<http://www.cnaf.infn.it/~sergio/datatag/glue/working/service>